Lecture notes on Control Systems applied to Ocean Energy Conversion

Duarte Valério

June 7, 2023

Contents

Introduction

These lecture notes introduce Control Theory to the students of the *Wave Energy* course. They have a threefold purpose:

- to provide the vocabulary and the basic notions needed to work together with control engineers;
- to exemplify some of the techniques needed when controlling devices to convert ocean energy;
- to pave the way for an in-depth study which will be based on specialised literature and take more time than that available in this short course.

This version of the lecture notes includes all the subjects covered until 2021. As the course has been reorganised, an abridged version is also available, covering only the subjects that can now be addressed. The full version is still provided, for the benefit of students who may need to deepen some particular subject.

Chapter 1

Basic concepts about signals and systems

1.1 Systems

System is the part of the Universe we want to study. *System*

A system made up of physical components may be called a **plant**. A system which is a combination of operations may be called a **process**. *Plant* *Process*

Example 1.1. A Tidal Energy Converter (TEC) is a plant. If we want to study the wave elevation at a certain onshore location as a function of the weather on the middle of the ocean, we will be studying a process.

The variables describing the characteristics of the system that we want to control are its **outputs**. *Outputs*

The variables on which the outputs depend are the system's **inputs**. *Inputs in the general sense*

The inputs of the system that cannot be modified are called **disturbances**. The inputs of the system we can modify are called **manipulated variables** or **inputs in the strict sense**. From now on, when referring to a variable as input, we mean that it is an input in the strict sense. *Disturbances* *Inputs in the strict sense*

Example 1.2. Consider a Wave Energy Converter (WEC) of the Oscillating Water Column (OWC) type, with a valve for pressure relief. To study this plant, we likely want to know, for each time instant, the electric power it is producing, the pressure inside the chamber, the rotation speed of the turbine, and the air mass flow through the turbine. So these will be its outputs. They depend on the incoming wave and on the position of the relief valve, which are the plant's inputs in the general sense. As we cannot modify the incoming wave, this will be a disturbance. As we can open and close the relief valve, its position is an input in the strict sense.

Remark 1.1. Control is by far easier when disturbances are neglectable. Remarkably, when extracting energy from the sea, they never are.

A system with only one input and only one output is a Single-Input, Single-Output (**SISO**) system. A system with more than one input and more than one output is a Multi-Input, Multi-Output (**MIMO**) system. It is of course possible *SISO system* *MIMO system*

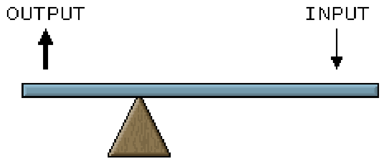

Figure 1.1: A linear system without dynamics (lever; source: Wikimedia Commons).

to have Single-Input, Multiple-Output (**SIMO**) systems, and Multiple-Input, Single-Output (**MISO**) systems. These are usually considered as particular cases of MIMO systems.

Example 1.3. The OWC of Example 1.2 is a MIMO plant. Another one is a car. When we are driving a car, we want to control its velocity and speed (the outputs). To do this, we can control the angle of the steering wheel, how far the gas pedal, the brake pedal and the clutch are operated, and which gear is engaged (the inputs; an automatic gearbox will mean less inputs). The wind gusts and the road conditions are disturbances. On the other hand, the lever in Figure 1.1 is a SISO system: if the extremities are at heights $x(t)$ and $y(t)$, and the first is actuated, then $y(t)$, the output, depends on position $x(t)$, the input, and nothing more.

Model

A system's **model** is the mathematical relation between its outputs, on the one hand, and its inputs and disturbances, on the other.

Linear system
Non-linear system

A system is **linear** if its exact model is linear, and **non-linear** if its exact model is non-linear. Of course, exact non-linear models can be approximated by linear models, and often are, to simplify calculations.

Example 1.4. The lever of Figure 1.1 is a linear plant, since, if its arm lengths are L_x and L_y for the extremities at heights $x(t)$ and $y(t)$ respectively,

$$y(t) = \frac{L_y}{L_x} x(t). \tag{1.1}$$

A Cardan joint (see Figure 1.2) connecting two rotating shafts, with a bent corresponding to angle β, is a non-linear plant, since a rotation of $\theta_1(t)$ in one shaft corresponds to a rotation of the other shaft given by

$$\theta_2(t) = \arctan \frac{\tan \theta_1(t)}{\cos \beta}. \tag{1.2}$$

A car is also an example of a non-linear plant, as any driver knows.

Dynamic system

A system has no dynamics if its outputs in a certain time instant do not depend on past values of the inputs or on past values of the disturbances. Otherwise, it is a **dynamic system**.

Example 1.5. Both mechanical systems in Figures 1.1 and 1.2 have no dynamics, since the output $y(t)$ only depends on the current value of the input $u(t)$.

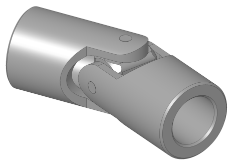

Figure 1.2: A non-linear mechanical system without dynamics (Cardan joint; source: Wikimedia Commons).

Past values of the input are irrelevant. The same happens with a faucet that delivers a flow rate $Q(t)$ given by

$$Q(t) = k_Q f(t) \tag{1.3}$$

where $f(t) \in [0, 1]$ is a variable that tells is if the faucet is open ($f(t) = 1$) or closed ($f(t) = 0$). But a faucet placed far from the point where the flow exits the pipe will deliver a flow given by

$$Q(t) = k_Q f(t - \tau) \tag{1.4}$$

This is an example of a dynamic plant, since its output at time instant t depends on a past value of $f(t)$. Here, τ is the time the water takes from the faucet to the exit of the pipe. And, again, a car is also an example of a dynamic system, as any driver knows.

1.2 Signals

A **signal** is a function of time or space that conveys information about a system. *Signal*

Example 1.6. One of the outputs of a WEC is the electric power it injects into the grid. This is a signal that depends on time. One of its disturbances is the wave it extracts energy from. Wave elevation is a signal that depends on both time and space.

Some signals can only take values in a discrete set; they are called **quantised signals**. Others can take values in a continuous set; they are called **analogical signals**. *Quantised signal* *Analogical signal*

Example 1.7. The rotation speed of a turbine is real valued; it takes values in a continuous set. So does the position of the break pedal of a car. The number of blades of the turbine is an integer number; it takes values in a discrete set. So does the shift engaged by the gearbox of a car.

Remark 1.2. In engineering, most signals (if not all) are **bounded**. For instance, the wave elevation at a certain point cannot be less than the depth of the sea, and cannot be arbitrarily large (a sea wave 1 km high, for instance, is a physical impossibility). Again, the rotation speed of a turbine, or the linear velocity of a shaft, or a voltage in a circuit, are always limited by physical constraints. Still, as these signals can assume values in a continuous set, there are infinite values they can assume. On the other hand, discrete signals, being limited, can only assume a limited number of values: a wave farm with a very large number of WECs can be conceived, but there is a physical limit for this too (probably well below the number of devices that would entirely cover all the oceans on Earth!).

Most signals are nowadays measured using digital equipment. Irrespective of the number of bits employed, there is a finite resolution implied, and so signals take only discrete values in practice. In other words, most signals are nowadays quantised.

Resolution

Precision

The **resolution** of a quantised signal is the difference between the consecutive discrete values that it may assume. In practice, this depends on how the signal is measured. The **precision** of a measurement is the range of values where the real value may be; in other words, it is the maximum error that can occur. Precision and resolution should not be confused.

Example 1.8. Figure 1.3 shows an input-output USB device. It can be used as an analog-to-digital (AD) converter. It reads signals in the -10 V to $+10$ V range using 12 bits. Consequently, it has an input resolution of $\frac{20}{2^{12}} = 4.88 \times 10^{-3}$ V. It can also be used as a digital-to-analog (DA) converter. It outputs signals in the 0 V to +5 V range using 12 bits. Consequently, it has an output resolution of $\frac{5}{2^{12}} = 1.22 \times 10^{-3}$ V.

Figure 1.3: National Instruments USB-6008, `http://www.ni.com/en-gb/support/model.usb-6008.html`.

Example 1.9. Analog measurements have resolution and precision too. Consider the weighing scale in Figure 1.4. Mass can be measured in the 5 kg to 100 kg range, with a resolution of 0.1 kg, and a precision of 0.1 kg. There is no finer resolution because the graduation has ten marks per kilogram. The precision is a result of the characteristics of the device.

Example 1.10. In Example 1.9 the resolution and the precision have the same value, but this is often not the case. Figure 1.5 shows luggage scales with a resolution of 1 g, but with a precision of 5 g or 10 g depending on the range where the measurement falls. (Varying precisions are found for some types of sensors, especially because of non-linearities.)

Remark 1.3. It makes sense to have a resolution equal to the precision, as in Example 1.9, in which case all figures of the measurement are certain; and it makes sense to have a resolution higher than the precision, as in Example 1.10,

Figure 1.4: Weighing scales once used in the Bełchatów coal mine, Poland.

Figure 1.5: Luggage scales.

in which case the last figure of the measurement is uncertain. It would make no sense to have a resolution more than 10 times larger than the precision, since in that case at least the last figure of the measurement would have no significance. It would also make no sense to have a resolution coarser than the precision, since the capacities of the sensor would be wasted.

Remark 1.4. The precision of a signal depends on all the elements of the measuring chain. The value shown at the display of the device in Figure 1.5 has a precision resulting from both the sensor used, and its particular precision, and the AD converter that the sensor's signal goes through, with its precision.

Remark 1.5. A naturally quantised signal can be measured with a coarser resolution. For instance, the population of a country has a resolution of one, but very often statistics give values rounded to thousands. This is because the uncertainty of the measured signal does not allow for finer resolutions, the last figures of which would have no significance.

Continuous signal
Discrete signal
Sampling time
Sampling frequency

Some signals take values for all time instants: they are said to be **continuous**. Others take values only at some time instants: they are said to be discrete in time, or, in short, **discrete**. The time interval between two consecutive values of a discrete signal is the **sampling time**. The sampling time may be variable (if it changes between different samples), or constant. In the later case, which makes mathematical treatment far more simple, the inverse of the sampling time is the **sampling frequency**.

Example 1.11. The air pressure inside the chamber of an OWC is a continuous signal: it takes a value for every time instant. The number of students attending the several classes of a course along the semester is a discrete signal: there is a value for each class, and the sampling time is the time between consecutive classes. The sampling time may be constant (if there is e.g. one lecture every Monday) or variable (if there are e.g. two lectures per week on Mondays and Tuesdays).

Most sensors nowadays measure the signal they are intended for only at some time instants, that is to say, with a given sampling time. In other words, nearly all measurements are discretised in time, just as they are quantised. If we want to store our data, and since data is normally recorded digitally, this makes all sense, as it would of course be impossible to record digitally a signal for *all* time instants. (An analogical record may sometimes be possible.)

Discretising a signal has to be done with care. If the sampling time is too big, we will miss many intermediate values that may be important. If it is too small, we will end up with many consecutive measurements that are either equal (because the signal does not change that fast) or where changes are irrelevant (because the measured value changes only due to some noise source).

Example 1.12. It is intuitive that, to study tides, we need not measure the sea level with a 1 ms sampling time (i.e. a 1000 Hz sampling frequency), as the period of tides is in the range of hours. It is also intuitive that, to study sea waves, we cannot measure the sea level once every minute, as we would miss most of the wave crests and troughs.

Nyquist theorem

While there is a theorem (the Nyquist theorem) about the lowest possible sampling time that can be used to sample a periodic signal, in practice a higher sampling frequency should be employed. The following rule of thumb is a fair indication of how the sampling time should be chosen: let ω_b be the highest frequency, in rad/s, we may be interested in studying. Then the sampling time T_s should verify

Rule of thumb for T_s

$$\frac{2\pi}{20\omega_b} \leq T_s \leq \frac{2\pi}{10\omega_b} \tag{1.5}$$

Bandwidth

(Frequency ω_b should more precisely be the system's bandwidth; we will mention this later on.) If $t_b = \frac{2\pi}{\omega_b}$ is the smallest time interval we are interested in studying, (1.5) is the same as

$$\frac{t_b}{10} \leq T_s \leq \frac{t_b}{20} \tag{1.6}$$

Example 1.13. If a tide has a period of 12 hours, then it should suffice to measure the sea level every 1.2 hours (70 minutes) at the least, or every 0.6 hours (36 minutes) at the most. As (1.5)–(1.6) are a rule of thumb, there would likely be no problem in measuring the sea level every half hour, or even every quarter of an hour.

Example 1.14. If we are interested in studying the waves at a shore where we know that during the entire year the wave spectra have neglectable content above 2 Hz = 12.57 rad/s, then a sampling time between 25 ms and 50 ms will be in order.

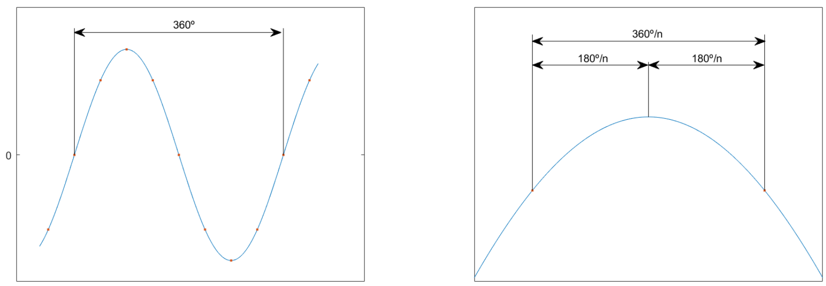

Figure 1.6: Sampled sinusoid. Left: in a very luck situation, sampling instants fall on zero crossings, crests and troughs. Right: in an equally unlucky situation, a crest is missed by as much as possible.

The lower value for the sampling time in rule 1.5 can be justified in the following way. When sampling a sinusoid, with some luck, sampling instants may fall on zero crossings and extreme values (crests and troughs), as seen in Figure 1.6. With an equal lack of luck, zero crossings and extreme values will be missed by as much as possible, as also shown in the Figure 1.6, for the case in which there are n sampling instants per period. If $n = 10$, this corresponds to an error of 18°. Since

$$\cos 18° = 0.95 \tag{1.7}$$

we see that, using 10 points per period, the amplitude of the sinusoid can be found from sampled data with an error of, at most, 5%. Decreasing the sampling time, lower errors in the measured amplitude will be obtained.

1.3 Models

There are basically two ways of modelling a system:

1. A model based upon **first principles** is a theoretical construction, resulting from the application of physical laws to the components of the plant. *First principles model*

2. A model based upon **experimental data** results from applying identification methods to data experimentally obtained with the plant. *Experimental model*

It is also possible to combine both these methods.

1.3.1 Models based upon first principles

These models can be obtained whenever the way the system works is known. They are the only possibility if the plant does not exist yet because it is still being designed, or if no experimental data is available. They may be quite hard to obtain if the system comprises many complicated interacting sub-parts. Simplifications can bring down the model to more manageable configurations, but its theoretical origin may mean that results will differ significantly from reality if parameters are wrongly estimated, if too many simplifications are assumed, or if many phenomena are neglected.

Example 1.15. A WEC consisting in a vertically heaving buoy of mass m can be modelled using Newton's law:

$$m\ddot{x(t)} = \sum F \tag{1.8}$$

Here $x(t)$ is the vertical position of the buoy (likely measured around its position for a calm sea). Finding expressions for all the forces involved — excitation force, radiation force, power take-off (PTO) force… — is not a trivial task, but good approximations can easily be found; in the end an added mass will turn up, as you know; and after some calculations something like this may be obtained:

$$(m + m_\infty)\ddot{x}(t) = F_e(t) + \int_{-\infty}^{t} h(t-\tau)\dot{x}(t)\,\mathrm{d}\tau - \rho g S x(t) + F_{\mathrm{PTO}}(t) \tag{1.9}$$

Dynamic systems can be modelled using **differential equations** if variables involved are continuous. This is the case of (1.8)–(1.9) above. If variables involved are discrete, **difference equations** are used instead (see examples below in section 1.3.5). Both models using differential equations and models using difference equations can be, as said in section 1.1, linear or non-linear, and SISO or MIMO. *Differential equations* *Difference equations*

In what follows we will assume linear models, and time-invariant parameters (**LTI models**). *LTI models*

Example 1.16. A plane consumes enormous amounts of fuel. Its mass changes significantly from take-off to landing. Any reasonable model of a plane will have to have a time-varying mass. But it is possible to study a plane, for a short period of time, using an LTI model, as the mass variation is neglectable in that case. WECs can have time-varying parameters due for instance to the effects of tides.

1.3.2 Continuous systems: Laplace transforms

Dynamic systems modelled using differential equations benefit enormously from the use of the Laplace transform, that allows turning a differential equation into an algebraic equation, far simpler to solve.

Definition of $\mathscr{L}$

The **Laplace transform** of a real-valued function $f(t)$ of positive real variable $t \in \mathbb{R}_0^+$ is given by

$$\mathscr{L}\left[f(t)\right] = \int_0^{+\infty} f(t) e^{-st}\, \mathrm{d}t \tag{1.10}$$

where $s \in \mathbb{C}$ is a complex variable, and e^{-st} is called the kernel of the integral transform (1.10). This transform exists if either

- $f(t)$ remains limited, or
- in the case $\lim\limits_{t\to+\infty} f(t) = \pm\infty$, if $f(t)$ diverges to infinity slower than the kernel e^{-st} goes to 0.

In either way, the integral in (1.10) exists. $\mathscr{L}\left[f(t)\right]$ is often denoted by $F(s)$.

Heaviside function

Example 1.17. Consider the Heaviside function,

$$H(t) = \begin{cases} 0, & \text{if } t < 0 \\ 1, & \text{if } t \geq 0 \end{cases} . \tag{1.11}$$

$\mathscr{L}\left[H(t)\right]$

Its Laplace transform is given by

$$\mathscr{L}\left[H(t)\right] = \int_0^{+\infty} H(t) e^{-st}\, \mathrm{d}t = \left[\frac{e^{-st}}{-s}\right]_0^{+\infty} = \frac{e^{-\infty}}{-s} - \frac{e^0}{-s} = \frac{1}{s}. \tag{1.12}$$

$\mathscr{L}\left[e^{-at}\right]$

Example 1.18. The Laplace transform of a negative exponential is

$$\mathscr{L}\left[e^{-at}\right] = \int_0^{+\infty} e^{-at} e^{-st}\, \mathrm{d}t = \left[\frac{e^{-(a+s)t}}{-a-s}\right]_0^{+\infty} = -\frac{e^{-\infty}}{s+a} - \left(-\frac{e^0}{s+a}\right) = \frac{1}{s+a}. \tag{1.13}$$

While Laplace transforms can be found from definition as in the examples above, in practice they are found from tables, such as the one in Table 1.1, which can also be used to find inverse Laplace transforms (i.e. finding the $f(t)$ corresponding to a given $\mathscr{L}\left[f(t)\right]$).

Laplace transforms are linear (the demonstration is straightforward) and have other properties, some of which are given in Table 1.2. One that is most

Table 1.1: Table of Laplace transforms

	$x(t)$	$X(s)$
1	$\delta(t)$	1
2	$H(t)$	$\frac{1}{s}$
3	t	$\frac{1}{s^2}$
4	t^2	$\frac{2}{s^3}$
5	e^{-at}	$\frac{1}{s+a}$
6	$1-e^{-at}$	$\frac{a}{s(s+a)}$
7	te^{-at}	$\frac{1}{(s+a)^2}$
8	$\sin(\omega t)$	$\frac{\omega}{s^2+\omega^2}$
9	$\cos(\omega t)$	$\frac{s}{s^2+\omega^2}$
10	$e^{-at}\sin(\omega t)$	$\frac{\omega}{(s+a)^2+\omega^2}$
11	$e^{-at}\cos(\omega t)$	$\frac{s+a}{(s+a)^2+\omega^2}$
12	$\frac{1}{b-a}\left(e^{-at}-e^{-bt}\right)$	$\frac{1}{(s+a)(s+b)}$
13	$\frac{1}{ab}\left(1+\frac{1}{a-b}\left(b\,e^{-at}-a\,e^{-bt}\right)\right)$	$\frac{1}{s(s+a)(s+b)}$
14	$\frac{\omega}{\Xi}\,e^{-\xi\omega t}\sin(\omega\Xi t)$	$\frac{\omega^2}{s^2+2\xi\omega s+\omega^2}$
15	$-\frac{1}{\Xi}\,e^{-\xi\omega t}\sin(\omega\Xi t-\phi)$	$\frac{s}{s^2+2\xi\omega s+\omega^2}$
16	$1-\frac{1}{\Xi}\,e^{-\xi\omega t}\sin(\omega\Xi t+\phi)$	$\frac{\omega^2}{s\left(s^2+2\xi\omega s+\omega^2\right)}$

In this table: $\Xi=\sqrt{1-\xi^2}$; $\phi=\arctan\frac{\Xi}{\xi}$

Table 1.2: Laplace transform properties

	$x(t)$	$X(s)$
1	$Ax_1(t) + Bx_2(t)$	$AX_1(s) + BX_2(s)$
2	$ax(at)$	$X\left(\frac{s}{a}\right)$
3	$e^{at}x(t)$	$X(s-a)$
4	$\begin{cases} x(t-a) & t > a \\ 0 & t < a \end{cases}$	$e^{-as}X(s)$
5	$\frac{dx(t)}{dt}$	$sX(s) - x(0)$
6	$\frac{d^2x(t)}{dt^2}$	$s^2X(s) - sx(0) - x'(0)$
7	$\frac{d^nx(t)}{dt^n}$	$s^nX(s) - s^{n-1}x(0) - \cdots - x^{(n-1)}(0)$
8	$-tx(t)$	$\frac{dX(s)}{ds}$
9	$t^2x(t)$	$\frac{d^2X(s)}{ds^2}$
10	$(-1)^nt^nx(t)$	$\frac{d^nX(s)}{ds^n}$
11	$\int_0^t x(u)\,du$	$\frac{1}{s}X(s)$
12	$\int_0^t \cdots \int_0^t x(u)\,du = \int_0^t \frac{(t-u)^{(n-1)}}{(n-1)!}x(u)\,du$	$\frac{1}{s^n}X(s)$
13	$x_1(t) * x_2(t) = \int_0^t x_1(u)\,x_2(t-u)\,du$	$X_1(s)\,X_2(s)$
14	$\frac{1}{t}x(t)$	$\int_s^\infty X(u)\,du$
15	$x(t) = x(t+T)$	$\frac{1}{1-e^{-sT}}\int_0^T e^{-su}X(u)\,du$
16	$x(0)$	$\lim_{s\to\infty} sX(s)$
17	$x(\infty) = \lim_{t\to\infty} x(t)$	$\lim_{s\to 0} sX(s)$

interesting can be obtained applying integration by parts to definition (1.10):

$$\mathscr{L}[f(t)] = \int_0^{+\infty} f(t)e^{-st}\,\mathrm{d}t = \left[f(t)\frac{e^{-st}}{-s}\right]_0^{+\infty} - \int_0^{+\infty} f'(t)\frac{e^{-st}}{-s}\,\mathrm{d}t =$$
$$= \lim_{t\to+\infty}\left(f(t)\frac{e^{-st}}{-s}\right) - f(0)\frac{e^0}{-s} + \frac{1}{s}\int_0^{+\infty} f'(t)e^{-st}\,\mathrm{d}t \tag{1.14}$$

The limit in the first term must be zero, or the Laplace transform would not exist. And the integral in the last term is $\mathscr{L}[f'(t)]$. Rearranging terms, we get

$$\mathscr{L}[f'(t)] = s\mathscr{L}[f(t)] - f(0) \tag{1.15}$$

where $f(0)$ is an **initial condition**. *Initial conditions*

1.3.3 Continuous systems: transfer functions

Consider a system with a dynamic behaviour described by

$$y(t) = 2u(t) - 3\dot{y}(t) \tag{1.16}$$

where $y(t)$ is the output, and $u(t)$ the input. We can apply Laplace transforms to the equation above, assuming that initial conditions are zero ($y(0) = 0$), and obtain

$$Y(s) = 2U(s) - 3sY(s) \Leftrightarrow Y(s) + 3sY(s) = 2U(s) \tag{1.17}$$

which can be rearranged into

$$\frac{Y(s)}{U(s)} = \frac{2}{1+3s} \tag{1.18}$$

This ratio of two polynomials in s is called a **transfer function** (relating the input $u(t)$ with the output $y(t)$, or, more exactly, the Laplace transform of the input $U(s)$ with the Laplace transform of the output $Y(s)$). Notice that to obtain the transfer function we had to assume **zero initial conditions**, otherwise we would not have arrived at a rational function of s. Transfer functions are often denoted by a capital letter such as $G(s)$.

(1.18) allows finding the output of the system very easily. Suppose that the input is $u(t) = 10H(t)$, where $H(s)$ is the Heaviside function. According to (1.12), and because the Laplace transform is linear, $U(s) = \frac{10}{s}$. Replacing this in (1.18) we get

$$Y(s) = \frac{20}{s(1+3s)} \tag{1.19}$$

and from Table 1.1, formula 6, we see that this corresponds to $y(t) = 20(1 - e^{-\frac{t}{3}})$. Check the calculations: they are easier than those that would have solved the problem employing (1.16) directly. The more complicated the differential equation and the expression of the input, the greater the advantage of solving only algebraic equations thanks to the Laplace transform.

Whenever you see a transfer function, remember that it is nothing more than a differential equation in disguise.

Example 1.19. (1.9) becomes (check the calculations)

$$X(s)\left[(m+m_\infty)s^2 - sH(s) + \rho g S\right] = F_e(s) + F_{\mathrm{PTO}}(s). \tag{1.20}$$

MIMO plants relate their several inputs and outputs with a matrix of transfer functions, such as

Transfer function ma

$$\begin{bmatrix} Y_1(s) \\ Y_2(s) \end{bmatrix} = \begin{bmatrix} \frac{1}{s^2+2s+1} & 0 & \frac{s-10}{s^2+2s+1} \\ \frac{s}{s^2+2s+1} & \frac{5}{s^2+2s+1} & \frac{-s+1}{s^2+2s+1} \end{bmatrix} \begin{bmatrix} U_1(s) \\ U_2(s) \\ U_3(s) \end{bmatrix} \tag{1.21}$$

1.3.4 Continuous systems: state-space representations

Consider a plant described by the differential equation

$$y(t) = u(t) - \dot{u}(t) - 2\dot{y}(t) - \ddot{y}(t) \tag{1.22}$$

State-space

This can be rearranged in the following form, called a **state-space representation**:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) \tag{1.23}$$

$$y(t) = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{1.24}$$

To prove that (1.23)–(1.24) is in fact the same as (1.22), it is easier to apply a Laplace transform (and drop the dependence on s for ease of notation):

$$sX_1 = X_2 \tag{1.25}$$

$$sX_2 = -X_1 - 2X_2 + U \tag{1.26}$$

$$Y = X_1 - X_2 \tag{1.27}$$

Replacing (1.25) in (1.26) we easily get

$$X_1 = \frac{U}{s^2+2s+1} \tag{1.28}$$

$$X_2 = \frac{sU}{s^2+2s+1} \tag{1.29}$$

$$Y = \frac{-s+1}{s^2+2s+1}U \tag{1.30}$$

States

and this is the transfer function corresponding to (1.22), as can be easily seen. The variables in vector $[x_1(t)\ x_2(t)]$ are the system's **states**.

A system with transfer function

$$G(s) = \frac{\sum_{k=0}^{n} b_k s^k}{s^n + \sum_{k=0}^{n-1} a_k s^k} = \frac{b_0 + b_1 s^1 + b_2 s^2 + \ldots + b_n s^n}{a_0 + a_1 s^1 + a_2 s^2 + \ldots + s^n} \tag{1.31}$$

Controllable canonical form

can be represented in state-space form in its so-called **controllable canonical form**

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{bmatrix} = \overbrace{\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{bmatrix}}^{\mathbf{A}} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{bmatrix} + \overbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}}^{\mathbf{B}} u(t) \tag{1.32}$$

$$y(t) = \underbrace{\begin{bmatrix} b_0 - a_0 b_n & b_1 - a_1 b_n & \dots & b_{n-1} - a_{n-1} b_n \end{bmatrix}}_{\mathbf{C}} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} + \underbrace{b_n}_{\mathbf{D}} u(t) \tag{1.33}$$

or in its so-called **observable canonical form**

Observable canonical form

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{bmatrix} = \overbrace{\begin{bmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{bmatrix}}^{\mathbf{A}} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{bmatrix} + \overbrace{\begin{bmatrix} b_0 - a_0 b_n \\ b_1 - a_1 b_n \\ \dots \\ b_{n-1} - a_{n-1} b_n \end{bmatrix}}^{\mathbf{B}} u(t) \tag{1.34}$$

$$y(t) = \underbrace{\begin{bmatrix} 0 & 0 & \dots & 0 & 1 \end{bmatrix}}_{\mathbf{C}} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{bmatrix} + \underbrace{b_n}_{\mathbf{D}} u(t) \tag{1.35}$$

There are in fact infinite possible representations of a system in state-space form, as the states of a system are not unique: any invertible linear transformation can be applied to a state-space representation, resulting in yet another state-space representation. The two above are just the most usual ones.

Infinite possible state vectors

The matrixes in the state-space representation are usually denoted by **A**, **B**, **C** and **D**, as indicated above. MIMO systems can easily be represented in state-space: they will have a vector of inputs, a vector of outputs, and bigger matrixes.

MIMO systems in state-space

1.3.5 Discrete systems: transfer functions and state-space representations

As mentioned in section 1.3.1, discrete systems are represented by difference equations, such as

$$y_k = \frac{1}{4}u_{k-1} + y_{k-1} - \frac{1}{4}y_{k-2} \tag{1.36}$$

Here u_k is the k-th sample of the input, and y_k is the k-th sample of the output. In other words,

$$y(k\,T_s) = \frac{1}{4}u((k-1)T_s) + y((k-1)T_s) - \frac{1}{4}y((k-2)T_s) \tag{1.37}$$

Delay operator z^{-1}

Difference equations are often rewritten using the **delay operator** z^{-1}, defined so that

$$z^{-1}y_k = y_{k-1} \tag{1.38}$$

$$z^{-2}y_k = y_{k-2} \tag{1.39}$$

$$z^{-3}y_k = y_{k-3} \tag{1.40}$$

and so on. For instance, (1.36)–(1.37) becomes

$$y = \frac{1}{4}z^{-1}u + z^{-1}y - \frac{1}{4}z^{-2}y \Leftrightarrow \frac{y}{u} = \frac{z^{-1}}{z^{-2} - 4z^{-1} + 4} \tag{1.41}$$

Discrete transfer function

Forward operator z

This last relation, a rational function of z^{-1}, is known as **discrete transfer function**. It can also be written with the **forward operator** z, defined so that

$$z\, y_k = y_{k+1} \tag{1.42}$$

$$z^2 y_k = y_{k+2} \tag{1.43}$$

$$z^3 y_k = y_{k+3} \tag{1.44}$$

and so on. For instance, (1.41) becomes

$$\frac{y}{u} = \frac{z}{1 - 4z + 4z^2}. \tag{1.45}$$

Just as signals can be discretised, so can continuous models given by transfer functions. An obvious way of doing so is using a first-order approximation of the derivative:

$$\dot{x}(k\,T_s) \approx \frac{x(k\,T_s) - x((k-1)T_s)}{T_s} \tag{1.46}$$

The Laplace of the left side is $sX(s)$. The right side corresponds to $\frac{x(1-z^{-1})}{T_s}$. Thus

$$s \approx \frac{1 - z^{-1}}{T_s} \Leftrightarrow s \approx \frac{z - 1}{z\,T_s} \Leftrightarrow z^{-1} \approx 1 - T_s s \Leftrightarrow z \approx \frac{1}{1 - T_s s} \tag{1.47}$$

For instance, replacing this, with a sampling time of $T_s = 1$ s in (1.22), we obtain (1.41) or (1.45). Other approximations exist; we will not study this matter further.

Discrete state-space

Discrete transfer functions can be put in state-space form too. Take for instance

$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix} u_k \tag{1.48}$$

$$y_k = \begin{bmatrix} \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} \tag{1.49}$$

where the states are x_1 and x_2. We can use the z-operator to rewrite this as

$$x_1 z = x_1 - \frac{1}{2}x_2 + \frac{1}{2}u \tag{1.50}$$

$$x_2 z = \frac{1}{2}x_1 \tag{1.51}$$

$$y = \frac{1}{2}x_1 \tag{1.52}$$

and from here it is not hard to show that (1.48)–(1.49) is the same as (1.41) or (1.45).

1.3.6 Models based upon both first principles and experimental data

Experimental data should, whenever available, be used to confirm, and if necessary modify, models based upon first principles. This often means that first principles are used to find a structure for a model (the orders of the derivatives in a differential equation, or the orders of the polynomials in a transfer function, or the size of matrixes in a state-space representation), and then the values of the parameters are found from experimental data: feeding the model the inputs measured, checking the results, and tuning the parameters until they are equal (or at least close) to measured outputs. This can sometimes be done using least squares; sometimes other optimisation methods, such as genetic algorithms, are resorted to.

If the outputs of experimental data cannot be made to agree with those of the model, when the inputs are the same, then another model must be obtained; this often happens just because too many simplifications were assumed when deriving the model from first principles. It may be possible to find, from experimental data itself, what modifications to model structure are needed.

1.3.7 Models based upon experimental data

Models based upon first principles can be called **white box models**, since the reason why the model has a particular structure is known. If experimental data requires changing the structure of the model, a physical interpretation of the new parameters may still be possible. The resulting model is often called a **grey box model**.

White box model

Grey box model

There are methods to find a model from experimental data that result in something that has no physical interpretation, neither is it expected to have. Still the resulting mathematical model fits the data available, providing the correct outputs for the inputs used in the experimental plant. Such models are called **black box models**, in the sense that we do not understand how they work. Such models are often **neural network models** (or NN models), or models based upon fuzzy logic (**fuzzy models**). We will not study these modelling techniques, but it is important to know that they exist.

Black box model
NN models
Fuzzy models

1.4 Exercises

1. Consider a heaving Wave Energy Converter with a dynamic behaviour given by

$$y(t) = 2 \times 10^{-6} F - 0.2\dot{y}(t) - 0.6\ddot{y}(t), \tag{1.53}$$

where y is the vertical position of the heaving element and F the force acting thereupon. Show that the WEC's transfer function is

$$\frac{Y(s)}{F(s)} = \frac{2 \times 10^{-6}}{0.6s^2 + 0.2s + 1}, \tag{1.54}$$

where $Y(s)$ and $F(s)$ are the Laplace transforms of the output and the input.

2. Find the differential equations represented in transfer function matrix (1.21).

3. Show that (1.53) is equivalent to the state-space representation

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & -\frac{5}{3} \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0.002 \\ 0 \end{bmatrix} F \tag{1.55}$$

$$y = \begin{bmatrix} 0 & \frac{5}{3} \times 10^{-3} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{1.56}$$

4. Show that (1.53) is equivalent as well to the state-space representation

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & 1 \\ -\frac{5}{3} & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{3} \times 10^{-5} \end{bmatrix} F \tag{1.57}$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{1.58}$$

   and verify that in fact a plant's state-space representation is not unique.

5. Find a state-space representation for each of the five non-null transfer functions in the matrix of (1.21).

6. Find a transfer function for the WEC with a dynamic behaviour given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & -\frac{5}{3} \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} F \tag{1.59}$$

$$y = \begin{bmatrix} -\frac{1}{6} \times 10^{-5} & \frac{1}{3} \times 10^{-5} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{1.60}$$

   and from the transfer function find the corresponding differential equation.

7. Find the differential equation corresponding to a WEC with a dynamic behaviour given by

$$\frac{Y(s)}{F(s)} = \frac{5 \times 10^{-6}}{s^2 + 3s + 2}, \tag{1.61}$$

   and find a state-space representation as well.

8. Discuss if a 100 ms sampling time is appropriate for the following systems, and propose a different value when it is not:

   (a) A system with a 100 rad/s bandwidth.
   (b) A system with a 5 rad/s bandwidth.
   (c) A system with a 0.3 rad/s bandwidth.
   (d) A tidal energy converter.
   (e) The controller of the opening of the valves in the engine of a car.
   (f) The controller of the opening of the valves in the diesel engine of a ship running at 80 rpm.

Chapter 2

Basic concepts about control

2.1 Supporting software

Many of the things mentioned in this chapter can be numerically verified using software packages such as **Matlab** or **Octave**. Matlab is a commercial product and you will need to buy it, to obtain a license (you may get one as a student of Técnico: check the DSI webpage), or to use a student version (with limited functionalities and valid for a limited time). Octave is a free software (see `https://www.gnu.org/software/octave/`) that implements many (though unfortunately not all) functionalities of Matlab. Finally, you may run Octave online at `https://octave-online.net/`, which is the simplest way to do so. If you do not have Matlab, and do not want to go through an Octave installation, just use this Octave Online website for the examples in this chapter.

Matlab
Octave
Octave Online

In what follows we will only use the simplest Octave (or Matlab) commands, so do not worry if you never used this software before. Just run the examples to check the results. They could all be found from Laplace transform tables, but, while using the Laplace transform is far simpler than solving the differential equations in other ways, it is still rather time consuming, and software to compute system responses numerically is even simpler to use.

2.2 Stability and robustness

A system is said to be **stable** if bounded inputs and bounded disturbances always result in bounded outputs. Systems for which such inputs and/or disturbances results in infinite outputs are called unstable. (Of course, in practice, the output never reaches an infinite value, but saturates instead; what is meant is that it would diverge to infinity if there were no limits to its value.)

Stability

Obviously, one of the objectives of a control system is to maintain stability, or achieve stability if the system is not stable.

Example 2.1. We saw in equation (1.13) that

$$\mathscr{L}^{-1}\left[\frac{1}{s+a}\right] = e^{-at}. \tag{2.1}$$

A transfer function with a pole (i.e. a root of the polynomial in the denominator) at $-a$ will always originate a time response with an exponential e^{-at}. This is true even if there are other poles, and is true irrespective of the system's input (that can be seen using a partial fraction expansion). If $a > 0$ the pole $-a$ is negative and the exponential tends to 0. But if $a < 0$ the pole is positive and the exponential diverges to infinity. Thus, any transfer function with at least one positive pole corresponds to an unstable system.

Stable poles

Unstable poles

Remark 2.1. If a transfer function has complex roots, it is the real part of the roots that originates an exponential that may either tend to 0 or diverge to infinity. Suppose for instance that $a \pm jb$ are poles. Then

$$\mathscr{L}^{-1}\left[\frac{1}{s+a+jb}\right] = e^{(-a-jb)t} = e^{-at}e^{-jbt} = e^{-at}\left(\cos bt - j\sin bt\right) \qquad (2.2)$$

$$\mathscr{L}^{-1}\left[\frac{1}{s+a-jb}\right] = e^{(-a+jb)t} = e^{-at}e^{jbt} = e^{-at}\left(\cos bt + j\sin bt\right). \qquad (2.3)$$

Stable poles are in the complex left half-plane

Notice that the imaginary parts cancel out, and the stability depends only on the real part of the poles.

Complex poles originate oscillations

Remark 2.2. Notice from (2.2)–(2.3) that the imaginary part of the poles originates a sinusoidal component. In other words, systems with transfer functions with complex poles have oscillatory outputs.

Complex poles appear in conjugate pairs

Remark 2.3. Remember then polynomials with real coefficients may have complex roots, but these always turn up in complex conjugate pairs.

Robustness to disturbances

Robustness to parameter uncertainty

Robustness of a control system is its ability of not being affected by undesirable disturbances (robustness to disturbances) or by changing parameters (robustness to parameter uncertainty).

2.3 Time responses

Systems are often studied and tested by the way the outputs evolve, in time, when the input is

- an impulse $\delta(t)$,
- a unit-step $H(t)$,
- a unit-slope ramp $r(t) = t$.

In practice, instead of an impulse, a pulse with a finite amplitude and a short duration must be used instead.

In practice, steps may be impossible, and a short but continuous variation of the input between two values is used instead. Also notice that 0 and 1 may not be acceptable values for the input, or an amplitude of 1 may be unfeasible. In the later case a different amplitude is used instead; the unit-step is used when possible just because calculations are simpler.

Example 2.2. It is obviously unfeasible to study the suspension of a car by applying a 1 m step to the tyres. The amplitude is too large: a step in the range between 0.01 m and 0.1 m is more reasonable. Studying the effects on the turbine of an OWC when there is a step of 1 Pa in the pressure of the chamber is similarly ridiculous, this time for the inverse reason.

In practice, a ramp can only be applied during a limited period of time, since infinite amplitudes of the input are always impossible. A unit-slope may also be unfeasible, and a different slope has then to be used instead.

2.3.1 First order systems

First order systems are described by differential equations with first order derivatives only, and, consequently, by transfer functions with first order polynomials only.

We will study the time responses of systems given by

$$G(s) = \frac{b}{s+a}. \tag{2.4}$$

This transfer function has a pole at $-a$, and does not have any zeroes.

Example 2.3. You can create transfer function $G_1(s) = \frac{15}{s+5}$ in Octave with the command

```
G1 = tf([15],[1 5])
```

(The coefficients of the numerator and the denominator go inside the square brackets.) Then plot the system's impulse response with the command

```
impulse(G1)
```

You can create transfer function $G_2(s) = \frac{20}{s+5}$ in the same way or with the commands

```
s = tf('s')
G2 = 20/(s+5)
```

Create transfer function $G_3(s) = \frac{25}{s+5}$ in either way and then plot their step responses using

```
step(G1,G2,G3)
```

See the result in Figure 2.1. Notice that after some time all responses tend to a constant value. They are said to be in **steady-state**. Before the steady-state is achieved, there is the **transient response**. Also notice that the steady-state response ends at $\frac{b}{a}$.

Steady-state response
Transient response

Example 2.4. Create now three transfer functions as follows:

```
Ga = tf(1,[1 1])
Gb = tf(2,[1 2])
Gc = tf(3,[1 3])
```

We can plot their step responses as we did above, or else defining a vector with time instants and a vector with the corresponding inputs:

```
t = 0 :  0.1 :  5;
input = ones(size(t));
```

(In either case, the semicolon at the end is optional; if you do not put it there, you will see the numbers in the vectors when you run the command. The first of these commands puts into variable `t` all numbers from 0 to 5, with a 0.1 step. The second puts in variable `input` a matrix of ones, with the same number of lines and columns of the matrix in variable `t`.) Then use command

```
lsim(Ga,Gb,Gc,input,t)
```

This is very useful when the input is not a unit step, since you can put any value you want in vector `input`. See the result in Figure 2.2. Notice that all transfer functions have a steady-state of $\frac{a}{b} = 1$, but they respond faster (slower) for larger (smaller) values of the pole a.

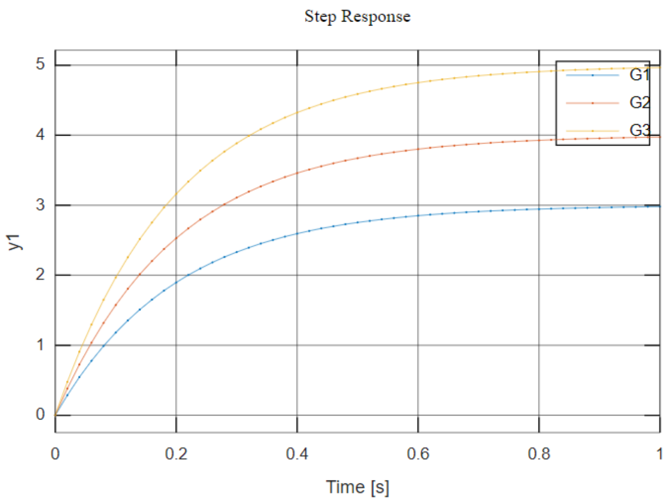

Figure 2.1: Unit-step responses of three first-order transfer functions with the same pole and different numerators.

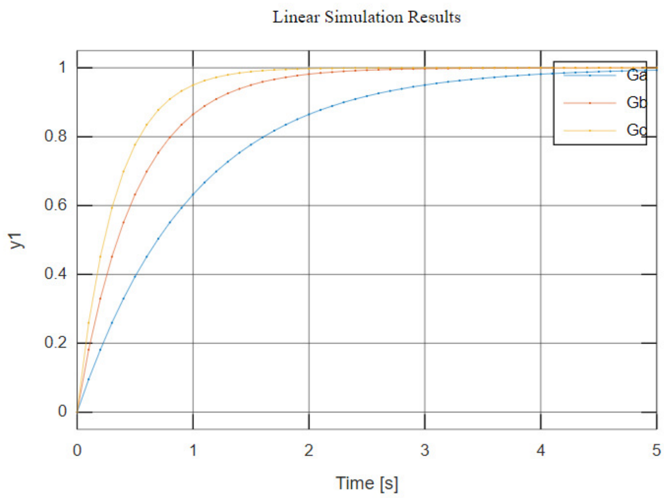

Figure 2.2: Unit-step responses of three first-order transfer functions with a different poles and numerators.

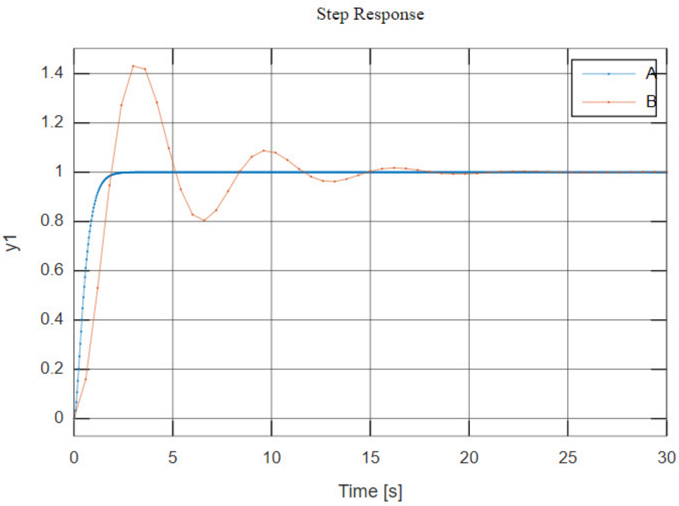

Figure 2.3: Unit-step responses of two second order transfer functions, one with real poles, and another with complex poles.

Example 2.5. Create this transfer function with a pole on the right half-plane:

```
s = tf('s')
G2 = 1/(s-1)
```

(You do not need to define `s` again if it already exists, but there is no harm in redefining it.) Plot its impulse and unit-step responses with commands `impulse` and `step`, as done above, and verify that, as expected, they diverge to infinity, since the plant in unstable.

2.3.2 Second order systems

Second order systems are described by differential equations with first and second order derivatives only, and, consequently, by transfer functions with second order polynomials only.

We will study the time responses of systems given by

$$G(s) = \frac{b}{s^2 + a_1 s + a_0}. \tag{2.5}$$

This transfer function has two poles and does not have zeroes.

Example 2.6. Create two transfer functions and plot their step responses:

```
s = tf('s')
A = 12/((s+3)*(s+4))
B = tf([1],[1 0.5 1])
step(A,B)
```

Notice that `A` has real poles and `B` has complex conjugate poles. As expected, there are no oscillations for real poles, but there are oscillations for complex poles, as shown in Figure 2.3.

2.3.3 Responses to sinusoids

LTI systems respond to a sinusoidal input with an output that, in steady state, is also a sinusoid with the same frequency.

Example 2.7. Let us define the following inputs:

```
t = 0 :  0.05 :  100;
input0_5 = sin(0.5*t);
input1 = sin(t);
input2 = sin(2*t);
```

Then we can plot time responses of `B` (from Example 2.6) as follows:

```
figure, lsim(B,input0_5,t)
hold, plot(t,input0_5)
```

(Command `figure` forces the creation of a new plot, while `hold` forces the plot to be superimposed upon the last one.)

```
figure, lsim(B,input1,t)
hold, plot(t,input1)
figure, lsim(B,input2,t)
hold, plot(t,input2)
axis([0 100 -1.2 1.2])
```

(This last command sets new limits to the area shown in the plot, since the automatic setting hides part of the input.) The results are shown in Figure 2.4. It is clear that, after a shorter or longer transient regime, the output becomes a sinusoid with the same frequency of the input.

2.4 Frequency responses

Frequency response

The above example of responses to sinusoidal inputs paves the way to the study of frequency responses. A system's **frequency response** is the study of how it responds, in steady-state, to a sinusoidal input. Since the output in steady-state is a sinusoid with the frequency of the input, two quantities suffice to describe it — both of them varying with the frequency:

Gain

- the **gain** is the ratio between the amplitude of the output sinusoid A_{output} and the amplitude of the input sinusoid A_{input};

Phase

- the **phase** is the difference in phase between the two sinusoids. If the output has its extreme values and zero crossings later than the input, the phase is negative; if it is the other way round, it is positive. The phase is, of course, defined up to an integer multiple of 360°: but continuous variations with frequency are assumed, beginning at the phase obtained for very low frequencies.

Example 2.8. Consider the frequency responses in Figure 2.4. For 0.5 rad/s, the gain is 1.26 (the amplitude of the output is larger than that of the input, which is 1) and the phase is $-19°$ (notice how the output is delayed in relation to the input). For 1 rad/s, the gain is 2 and the phase is $-90°$ (notice how the output crosses zero as the input is already at a peak or at a through). For 2 rad/s, the gain is 0.32 (the output is smaller than the input) and the phase is $-162°$ (notice how input and output are almost in phase opposition, i.e. when one has a peak the other has a through and vice-versa).

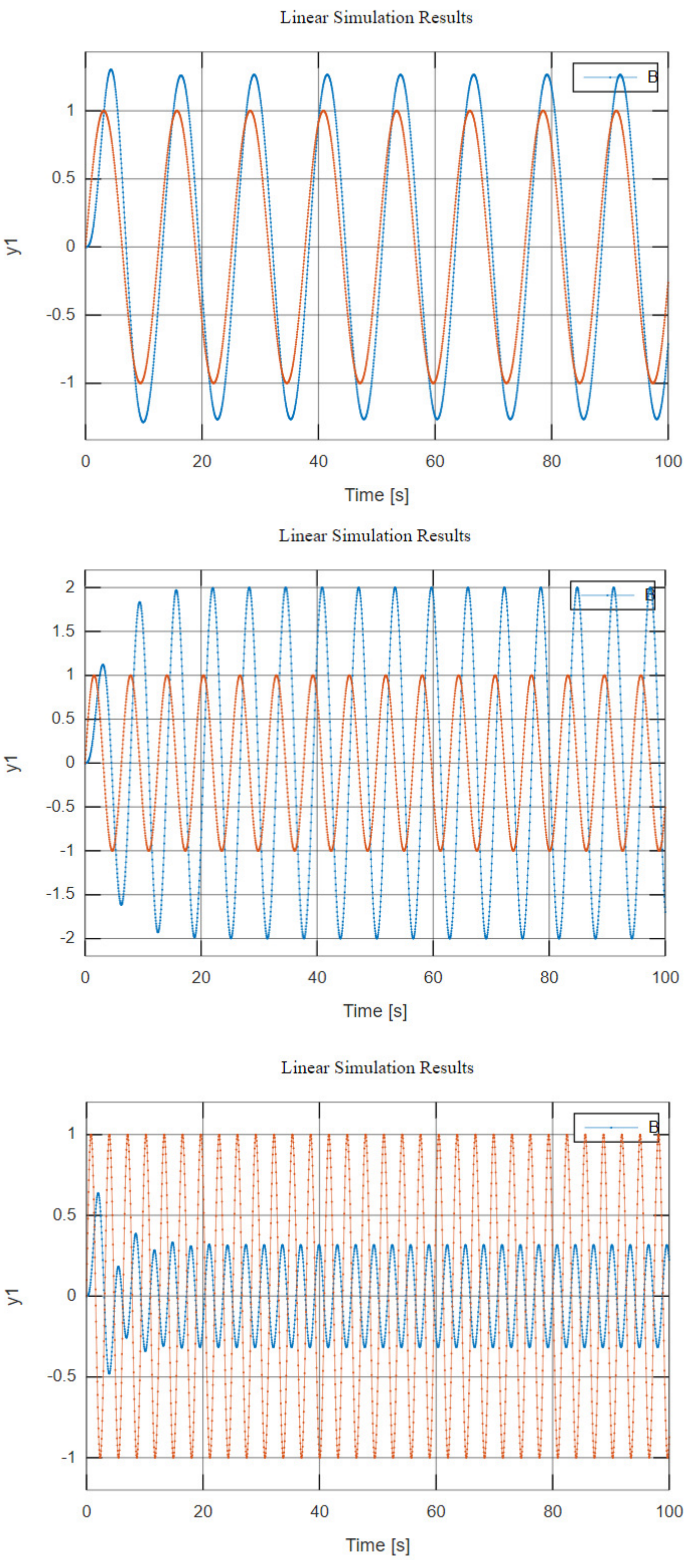

Figure 2.4: Responses of $B(s) = \frac{1}{s^2+0.5s+1}$ to sinusoids with 0.5 rad/s (top), 1 rad/s (centre) and 2 rad/s (bottom)

Decibel scale for gains

Gains are often given not in absolute value, but in **decibel**; i.e. instead of the ratio $\frac{A_{\text{output}}}{A_{\text{input}}}$, the value $20 \log_{10} \frac{A_{\text{output}}}{A_{\text{input}}}$ dB is given instead.

Remark 2.4. Notice that gains in absolute value are in the $]0, +\infty[$ range:

- a value of 1 means that the output and the input have the same amplitude,
- a value larger than 1 means that the output has an amplitude larger than that of the input,
- a value smaller than 1 means that the output has an amplitude smaller than that of the output.

Gains in decibel are in the $]-\infty, +\infty[$ range:

- a value of 0 dB means that the output and the input have the same amplitude,
- *Positive gains in dB* — a value larger than 0 dB means that the output has an amplitude larger than that of the input,
- *Negative gains in dB* — a value smaller than 0 dB means that the output has an amplitude smaller than that of the output.

Bode diagram

A **Bode diagram** is a graphical representation of the frequency response of a system, as a function of frequency. It actually comprises two plots:

- *Gain plot* — a top plot, where the gain in dB is shown as a function of the frequency, with a semilogarithmic scale on the x-axis;
- *Phase plot* — a bottom plot, where the phase in degrees is shown as a function of the frequency, also with a semilogarithmic scale on the x-axis.

Example 2.9. To plot the Bode diagram of `B` from Examples 2.6 and 2.7, we give the command

```
bode(B)
```

The result is shown in Figure 2.5. Notice how the gain for 2 rad/s is in fact $20 \log_{10} 0.32 = -9.9$ dB, the gain for 1 rad/s is in fact $20 \log_{10} 2 = 6.0$ dB, and the gain for 0.5 rad/s is in fact $20 \log_{10} 1.26 = 2.0$ dB. Also notice how phases have the expected values.

Example 2.10. Figure 2.5 shows that at 100 rad/s the gain is -80 dB (in absolute value, $10^{-80/20} = 1 \times 10^{-4}$) and the phase is $-180° = -\pi$ rad. So we know that if the input of $B(s)$ is

$$y(t) = \sin(100t) \tag{2.6}$$

the output will be

$$u(t) = 1 \times 10^{-4} \sin(100t - \pi). \tag{2.7}$$

Figure 2.5 also shows that at 0.01 rad/s the gain is 0 dB (in absolute value, 1) and the phase is 0°. So we know that if the input of $B(s)$ is

$$y(t) = \sin(0.01t) \tag{2.8}$$

the output will be

$$u(t) = \sin(0.01t) \tag{2.9}$$

as well.

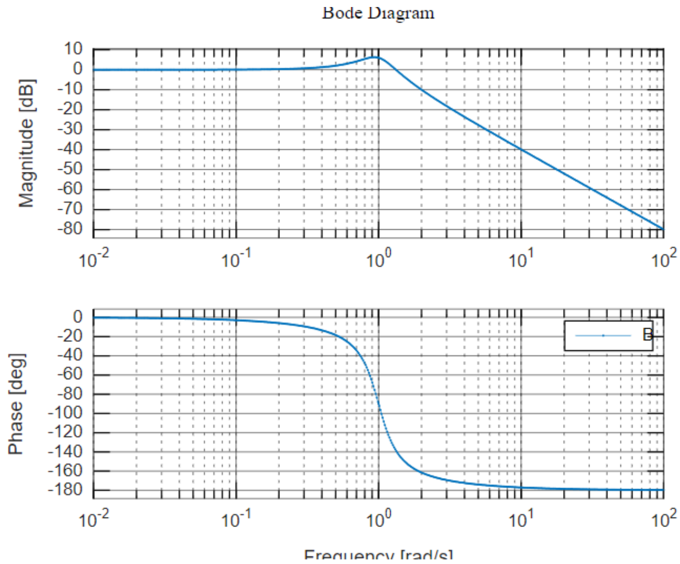

Figure 2.5: Bode diagram of $B(s) = \frac{1}{s^2+0.5s+1}$.

Example 2.11. Because $B(s)$ is linear, if the input is

$$y(t) = 12.345\sin(100t) \tag{2.10}$$

the output will be

$$u(t) = 12.345 \times 10^{-4}\sin(100t - \pi). \tag{2.11}$$

And, again because $B(s)$ is linear, if the input is

$$y(t) = 0.1\sin\left(0.01t + \frac{\pi}{6}\right) + \sin\left(0.5t - \frac{3\pi}{7}\right) + 3.5\sin\left(t - \frac{11\pi}{12}\right) + 0.2\sin\left(2t + \frac{2\pi}{5}\right) \tag{2.12}$$

the output will be

$$y(t) = 0.1\sin\left(0.01t + \frac{\pi}{6}\right) + 1.26\sin\left(0.5t - \frac{3\pi}{7} - \frac{19\pi}{180}\right) + 7\sin\left(t - \frac{11\pi}{12} - \frac{\pi}{2}\right) + 0.64\sin\left(2t + \frac{2\pi}{5} - \frac{162\pi}{180}\right). \tag{2.13}$$

In Naval Engineering the gain is often called **Response Amplitude Operator** (RAO) and represented in absolute value as a function of the frequency or the period (in linear scale). Notice that the RAO alone does not suffice to find the response to a sinusoidal input, since the information on the phase is missing. If the input is not a sinusoid, but a periodic signal, which can thus be decomposed into a sum of sinusoids (this is the case of sea waves), knowing both gain and phase is necessary to accurately find the output. *RAO*

Example 2.12. The RAO of $B(s)$ from Figure 2.5 is shown in Figure 2.6.

Instead of the Bode diagram — by far the most usual —, frequency responses may be shown in alternative diagrams, such as the **Nyquist diagram** or the **Nichols diagram**, which we will not study. *Nyquist diagram* *Nichols diagram*

Dynamic systems are called:

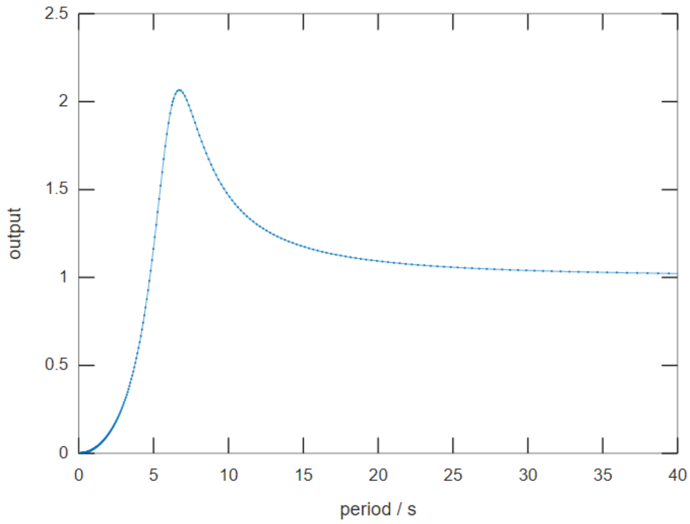

Figure 2.6: RAO of $B(s) = \frac{1}{s^2+0.5s+1}$.

Low pass filter

- **low-pass filters**, if they attenuate inputs at high frequencies, but not at low frequencies;

High pass filter

- **high-pass filters**, if they attenuate inputs at low frequencies, but not at high frequencies;

Band pass filter

- **band-pass filters**, if they attenuate inputs at both low and high frequencies, but not at some range of frequencies between.

Example 2.13. See Figure 2.7 for the Bode diagrams of:

- low-pass filter $G_l(s) = \dfrac{10}{s+10}$;
- high-pass filter $G_h(s) = \dfrac{s}{s+1}$;
- band-pass filter $G_b(s) = \dfrac{20s}{(s+0.5)(s+20)}$.

Bandwidth as range of frequencies

A system's **bandwidth** is the range of frequencies where its gain is above a certain threshold, important for a particular application. For reasons we need not enter into, a threshold of -3 dB is often used. With this criterion, outside its bandwidth, a system will attenuate its input by $10^{-3/20} = 0.71$ or more.

Bandwidth as the upper limit of a range of frequencies

For low pass filters, only the upper limit of the bandwidth is given. It is in this sense that the bandwidth in (1.5) should be understood.

Finally, here is a result, stated for reference purposes only, and that we will not prove: at frequency ω, transfer function $G(s)$ has a gain given by $|G(j\omega)|$, and a phase given by $\arg[G(j\omega)]$. This allows plotting Bode diagrams (or other graphical representations of the frequency response) rather easily, but you will not be asked to do so.

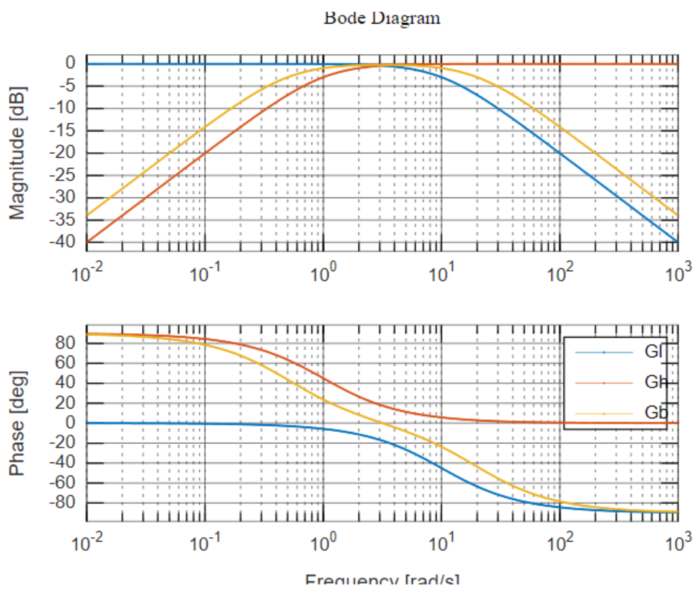

Figure 2.7: Bode diagrams of a low-pass filter, a high-pass filter, and a band-pass filter.

2.5 Open-loop control and closed-loop control

The two simplest configurations of control systems are open-loop control and closed-loop control.

In **open-loop control**, a control action is applied, and the result in the output is not verified. If the output is measured, this measurement is not used to correct the control action if there is some deviation from the desired value. In other words, there is no feedback of the output. *Open-loop control*

In **closed-loop control**, the value of the output is compared with the desired reference, and the **error** between the reference and the output is fed to the controller. The control action depends on this error. In other words, there is feedback of the output. *Closed-loop control* *Closed-loop error*

These two control strategies can be represented in **block diagrams** as seen in Figure 2.8, where: *Block diagrams*

- $G(s)$ is the transfer function of the system to be controlled;
- $C(s)$ is the transfer function of the controller;
- $U(s)$ is (the Laplace transform of) the control action provided by $C(s)$, and the input of $G(s)$;
- $Y(s)$ is (the Laplace transform of) the output of $G(s)$, i.e. the variable we want to control;
- $R(s)$ is (the Laplace transform of) the reference for $Y(s)$ to follow; *Reference*
- $E(s)$ is (the Laplace transform of) the closed-loop error.

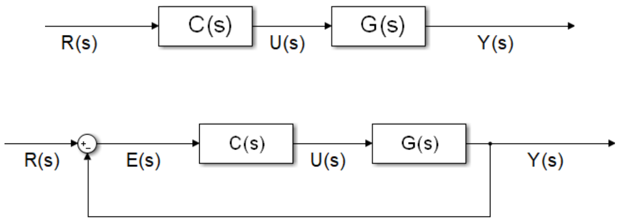

Figure 2.8: Block diagrams of open-loop control (top) and closed-loop control (bottom).

Since a block corresponds to a transfer function, and the incoming arrow corresponds to the Laplace transform of the input, the outcoming arrow corresponds to the Laplace transform of the output. So, thanks again to Laplace transforms, block diagram algebra is very simple, and shows that

- when using open-loop control,

$$Y(s) = C(s)G(s)R(s) \Leftrightarrow \frac{Y(s)}{R(s)} = C(s)G(s); \tag{2.14}$$

Closed-loop transfer function

- when using closed-loop control,

$$\begin{aligned} &Y(s) = C(s)G(s)E(s) \Leftrightarrow Y(s) = C(s)G(s)(R(s) - Y(s)) \Leftrightarrow \\ &Y(s) + C(s)G(s)Y(s) = C(s)G(s)R(s) \Leftrightarrow \frac{Y(s)}{R(s)} = \frac{C(s)G(s)}{1 + C(s)G(s)} \end{aligned} \tag{2.15}$$

Open-loop control only makes sense when the system is very well known, so that the control action needed in each moment can be determined precisely, without need to check the result.

Example 2.14. Open-loop control can be applied in TECs, for instance, because tides are well-known in what concerns both their amplitude and the hours at which they take place; or in a WEC, in a control system that copes with tides. It is very hard to apply open-loop control in the presence of signals hard to predict such as waves; even when there is a very good upstream measurement of the wave it is often better to apply closed-loop control, checking if the control action is achieving or not its purpose.

On-off control

Example 2.15. Open-loop control can be applied to achieve **on-off control**, in which the output can only assume two values. That is for instance the case of a valve that should either be open or closed, actuated by a solenoid: a control action is applied that will bring the valve to one of its two saturation limits, and there may be no need to check that this happened.

Closed-loop control has a corrective action and does not require a knowledge of the system as exact as that which is necessary to apply open-loop control, though some knowledge is required lest the controller should be badly designed,

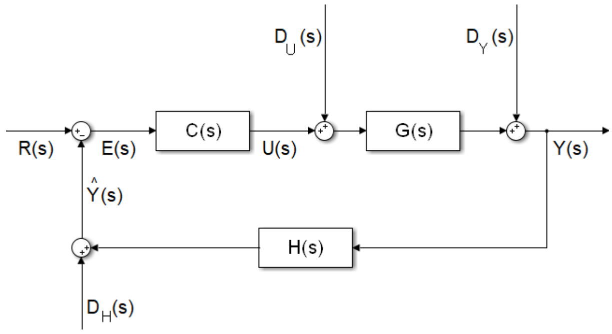

Figure 2.9: Block diagram of a closed-loop with disturbances.

preventing the closed-loop from achieving the desired performance, or even turning unstable. Notice that closed-loop control only reacts to deviations of the reference, and thus lacks any preventive or predictive action.

A more complete and realistic block diagram for closed-loop control is shown in Figure 2.9, where

- $D_U(s)$ is (the Laplace transform of) a disturbance affecting the control action;
- $D_Y(s)$ is (the Laplace transform of) a disturbance affecting the output;
- $D_H(s)$ is (the Laplace transform of) a disturbance affecting the measured output;
- $H(s)$ is the transfer function of the sensor that measures the output $Y(s)$;
- $\hat{Y}(s)$ is (the Laplace transform of) the measured value of the output.

Notice that an ideal sensor verifies $H(s) = 1$ and thus, if there is no sensor disturbance, $\hat{Y}(s) = Y(s)$.

2.6 Closed-loop controllers and how to design them

Closed-loop controllers can have many forms, among which the following are remarkable for their extended use:

- **Proportional controllers**: *Proportional control*

  $$C(s) = K_p \in \mathbb{R}, \tag{2.16}$$

  and so

  $$u(t) = K_p e(t). \tag{2.17}$$

  In this way, the larger (smaller) the error, the larger (smaller) the control action.

- **Proportional–integral (PI) controllers**: *PI control*

$$C(s) = K_p + \frac{K_i}{s},\ K_i \in \mathbb{R}, \tag{2.18}$$

and so

$$u(t) = K_p e(t) + \int_0^t e(t)\,\mathrm{d}t. \tag{2.19}$$

The control action will have a component proportional to the integral of the error, and so, if the error does not go to zero fast, the control action will increase to try to compensate this non-vanishing error.

PD control

- **Proportional–derivative (PD) controllers**:

$$C(s) = K_p + K_d s,\ K_d \in \mathbb{R}, \tag{2.20}$$

and so

$$u(t) = K_p e(t) + K_d \frac{\mathrm{d}e(t)}{\mathrm{d}t}. \tag{2.21}$$

The control action will have a component proportional to the derivative of the error, and so, if the error increases abruptly, the control action will increase to try to compensate this increasing error.

PID control

- **Proportional–integral–derivative (PID) controllers**:

$$C(s) = K_p + \frac{K_i}{s} + K_d s, \tag{2.22}$$

and so

$$u(t) = K_p e(t) + \int_0^t e(t)\,\mathrm{d}t + K_d \frac{\mathrm{d}e(t)}{\mathrm{d}t}. \tag{2.23}$$

This controller has three components, as its name says.

While controllers can have other forms, and have many poles and zeroes, the expressions above suffice to able to solve a great variety of control problems.

There are several different techniques to design controllers, among which the following deserve to be mentioned:

- There are several analytical techniques to decide which poles and zeroes the controller transfer function must have to achieve specifications related to time responses, to frequencies responses, or both. We will not study any, save for Internal Model Control, addressed below.

Ziegler-Nichols PID tuning rule for an S-shaped step response

- **Tuning rules** allow calculating controller parameters from a minimum of information about a plant. Among the many existing rules, the **Ziegler-Nichols tuning rules** for PIDs are particularly widespread; the first rule is applied to systems with an S-shaped unit-step response, as shown in Figure 2.10 (if the input step has an amplitude A, then the response will tend to KA). From the values in the Figure, the following controllers can be obtained:

$$C_P(s) = \frac{T}{LK} \tag{2.24}$$

$$C_{PI}(s) = \frac{0.9T}{LK}\left(1 + \frac{0.3}{Ls}\right) \tag{2.25}$$

$$C_{PID}(s) = \frac{1.2T}{LK}\left(1 + \frac{1}{2Ls} + 0.5Ls\right) \tag{2.26}$$

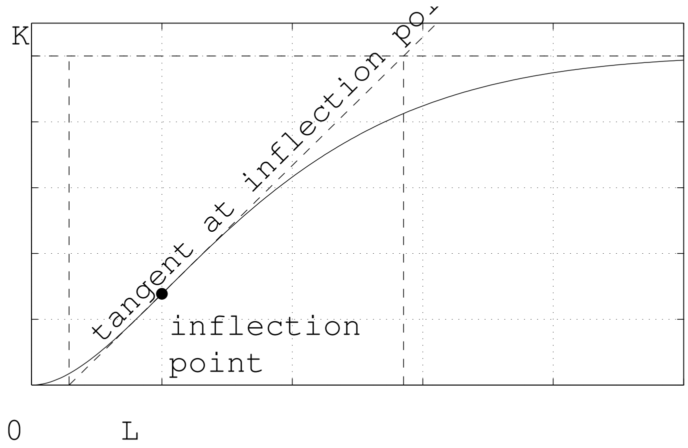

Figure 2.10: S-shaped step response.

Many other rules can be found in the literature. None can be expected to provide results as good as those of an analytical method.

- **Optimal control** consists in defining an **objective function** that should be minimised, usually penalising deviations of the output from the reference, big control actions, and big changes in the control actions (to spare the actuators); controller parameters are then found minimising this objective function. Analytical techniques can be found for LTI systems and quadratic objective functions, and for other well-behaved cases; these control techniques, based upon state-space representations, were among those known in the 1960's as Modern Control. More complicated situations may require other optimisation methods.

- **Predictive control** optimises an objective function as well, this time using a model of the system to predict how different particular control actions will affect performance during a period of time (the **prediction horizon**), finding the best option with a numerical search procedure (often based upon the **branch and bound** algorithm), implementing the best control for a period of time (the **control horizon**) shorter than the prediction horizon (because short-time predictions are more reliable than long-time predictions), and repeating the process (resulting in a **receding horizon** control).

- **Switching control** control consists in having different controllers to be applied in different situations. This has been used to control WECs under different sea states. Switching between controllers has to be done with care, since an abrupt change in controller can easily result in a big change of the control action, which may render the system unstable.

- **Non-linear control** is applied to plants with significant non-linearities. There are several techniques to do so, such as feedback linearisation or sliding-mode control; we will not study any details.

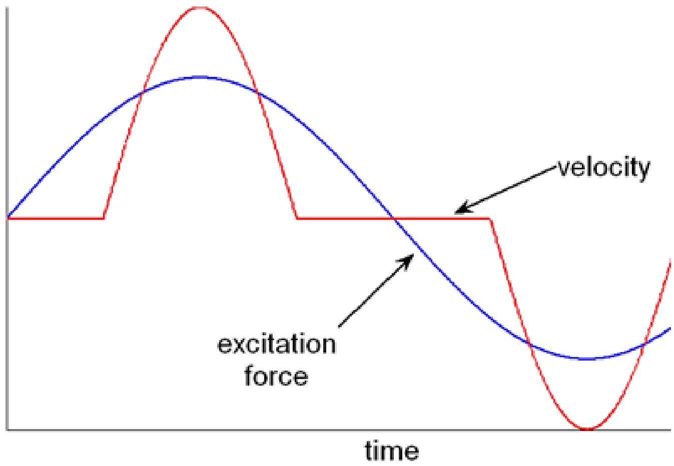

Figure 2.11: Objective of latching control for a heaving point absorber.

- **On-off control** has already been mentioned; it is in fact a very basic non-linear control technique based upon the saturation of the control variable.

Latching

- **Latching** is a non-linear control technique that consists in stopping an oscillating variable when it reaches a maximum or a minimum, to release when it is reckoned that it will be in phase with another variable. This has been extensively used with heaving WECs, which are latched when they stop either at the top or at the bottom of the movement (see Figure 2.11), so that the velocity will be in phase with the excitation force (so as to maximise the extracted power). As sea-waves are irregular, finding the instant to unlatch the WEC is not trivial, and since the mass that has to latched is usually significant, huge forces may be involved. This type of control increases the abrupt variations with time of the power to be injected into the grid; this is an undesirable consequence, that may be mitigated if there is a wave farm with many WECs, with the power variations of the different devices out of phase with each other. Latching control is similar to what happens in an OWC with a relief valve, that may as well be used to try to put the variations of pressure in phase with the excitation force.

2.7 Variations on closed-loop control

Cascade control

Many systems are made up of different parts in sequence, some faster than others. In that case, rather than trying to control the entire system at once, it is preferable to control separately the slower and the faster parts. This is the idea of a particular type of closed-loop control called **cascade control** (or master-slave control), with a block diagram shown in Figure 2.12, where:

- $G_1(s)$ is the faster part of the system,
- $G_2(s)$ is the slower part of the system.

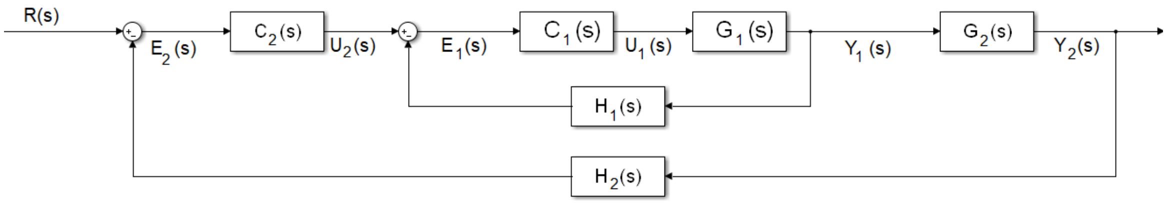

Figure 2.12: Block diagram of cascade control.

The remaining signals and blocks are self-evident. Notice that there must be two sensors, but there is still only one actuator, delivering control action $U_1(s)$ ($U_2(s)$ is just a signal that will be used as reference by controller $C_1(s)$). It is possible to have three, or more, nested loops in this way.

Example 2.16. Consider an OWC controlled by a valve that can assume any position between 0 (closed) and 1 (open). The input of the system is the tension applied to the actuator of the valve; the valve's position is affected by disturbances (e.g. a time-varying suction caused by the flow). Rather than designing only one controller that simultaneously handles the OWC and the valve, it is better to design first a controller to put the valve in the position desired. The dynamics of the valve will be $G_1(s)$; the controller (likely a PI or a PID) will be $C_1(s)$; a sensor $H_1(s)$ that measures the position of the valve $Y_1(s)$ is needed; the control action $U_1(s)$ is the tension that controls the valve. $G_2(s)$ will correspond to the dynamics of the OWC; $C_2(s)$ will only have to take care of *that*. If the inner closed-loop comprising $C_1(s)$, $G_1(s)$ and $H_1(s)$ is fast enough, the valve will get to a desired position much faster than the OWC reacts to a change: so $C_2(s)$ may even assume that the entire inner-loop has a transfer function of 1 (i.e. responds so fast that it is for practical purposes instantaneous). Even if this is not the case, at least the dynamics of that part of the system will be known and fixed in advance, so designing $C_2(s)$ will be easier. Figure 2.12 does not show disturbances, but they will exist: there are those affecting the valve, and there are of course the waves.

Internal Model Control (IMC) may be considered another variation of closed-loop control, that can be used when there is a good model of the system and a good inverse model of the system (i.e. a model that receives the system's output $y(t)$ as input, and delivers the system's input $u(t)$ as output). It is shown in Figure 2.13, where *IMC*

- $G(s)$ is the actual system to control,
- $G^*(s)$ is the model of the plant,
- $G^{-1}(s)$ is the inverse model of the plant.

Notice that if the model is perfect ($G^*(s) = G(s)$) then the error is

$$E(s) = R(s) - \Big(Y(s) - \hat{Y}(s)\Big) = R(s) - D(s) - G(s)U(s) + G^*(s)U(s) = R(s) - D(s). \tag{2.27}$$

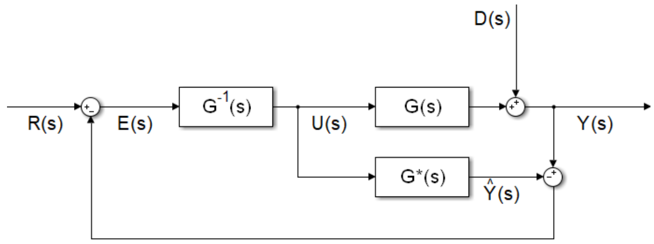

Figure 2.13: Block diagram of IMC.

If the inverse model is perfect, then

$$Y(s) = D(s) + G(s)G^{-1}(s)E(s) = D(s) + R(s) - D(s) = R(s) \tag{2.28}$$

In other words, with perfect models, IMC achieves perfect disturbance rejection, i.e. perfect robustness to disturbances. In practice, models are never perfect, but IMC often works well enough if models are good. IMC is suitable when using black-box modelling techniques. Notice that IMC is equivalent to a usual closed-loop (Figure 2.8) if

$$C(s) = \frac{G^{-1}(s)}{1 - G^{-1}(s)G^{*}(s)}. \tag{2.29}$$

Delay systems

Systems with a **delay** do not respond immediately to an input, but only after a while. Table 1.2, line 4, shows that the transfer function of a delay of θ seconds is the non-linear term $e^{-\theta s}$.

Example 2.17. Figure 2.14 shows the unit-step response of two transfer functions:

$$A(s) = \frac{8}{s^2 + 3s + 2} \tag{2.30}$$

$$B(s) = \frac{8}{s^2 + 3s + 2} e^{-2s} \tag{2.31}$$

Unfortunately, Octave does not deal with delay systems; Matlab does, and the figure was obtained with the following commands:

```
s = tf('s')
A = 8/((s+1)*(s+2))
B = A * exp(-2*s)
step(A,B)
legend('A','B')
```

Effects of delays in closed-loop

Delays are one of the biggest nuisances in control, as they easily cause closed-loop control systems to become unstable. The reason for this is intuitive: if a control action is applied, but there is no immediate response, the controller may easily increase the control action, to try to elicit some reaction from the system. The control action will then be so big that, when the system finally responds, it overshoots, and then the controller may easily try to solve this by decreasing too much the control action, resulting in an overshoot in the other direction, and so on. Delays are thus responsible not only for slower responses, but also for oscillations, and eventually unstable closed-loops.

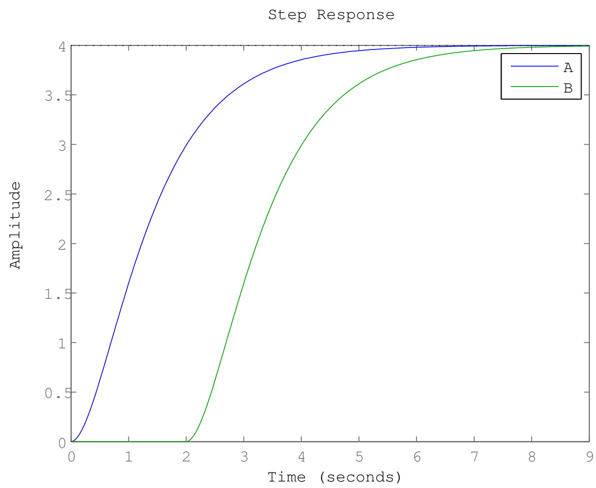

Figure 2.14: Unit-step responses of a system without delay and of a system with delay.

Example 2.18. You likely have had a similar experience when trying to take a shower in a bathroom you do not know, and where the water takes longer to reach the shower than you are used to. You will likely have turned too much the faucet of hot water, and then got burnt; then as reaction you probably turned the faucets so that water went too cold. In such cases it takes some time to get to know the system well enough to obtain a comfortable temperature.

What is needed is a controller that waits for the delay and takes it into account when determining the control action. This is possible with a variation of IMC called **Smith predictor**, shown in Figure 2.15, where *Smith predictor*

- $G(s)e^{-\theta s}$ is a system with a delay;
- $G^*(s)$ is a model of the system without delay;
- $\hat{\theta}$ is an estimate of the system's delay.

If $G^*(s)$ and $\hat{\theta}$ are exact (i.e. if the model is perfect), then $Y(s) = \hat{Y}(s)$, and thus what is being fed back is

$$G^*(s)U(s) = G^*(s)C(s)E(s) = G(s)C(s)E(s). \tag{2.32}$$

Consequently, the error is given by

$$E(s) = R(s) - G(s)C(s)E(s) \Leftrightarrow E(s)\left(1 + G(s)C(s)\right) = R(s) \Leftrightarrow E(s) = \frac{R(s)}{1 + G(s)C(s)}, \tag{2.33}$$

and since $Y(s) = G(s)e^{-\theta s}C(s)E(s)$ then

$$Y(s) = \frac{G(s)e^{-\theta s}C(s)R(s)}{1 + G(s)C(s)} \Leftrightarrow \frac{Y(s)}{R(s)} = \frac{G(s)e^{-\theta s}C(s)}{1 + G(s)C(s)} \tag{2.34}$$

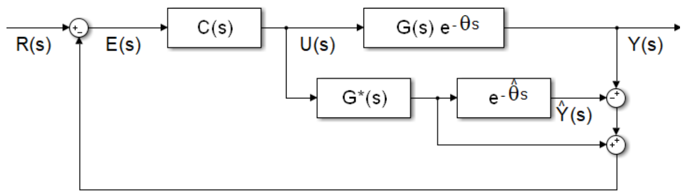

Figure 2.15: Smith predictor.

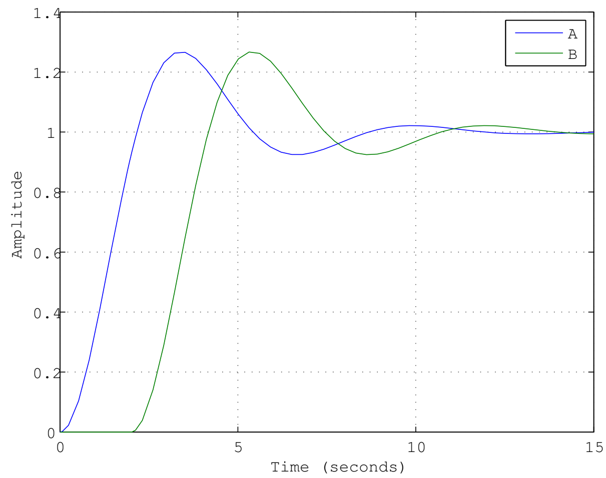

Figure 2.16: Unit-step responses of $A(s)$ controlled with $C(s)$ in closed-loop, and of $B(s)$ controlled with $C(s)$ and a Smith predictor.

Notice the difference between this and the result of a usual closed-loop:

$$\frac{Y(s)}{R(s)} = \frac{G(s)e^{-\theta s}C(s)}{1 + G(s)e^{-\theta s}C(s)} \tag{2.35}$$

Example 2.19. Figure 2.16 shows the unit-step response of $A(s)$, given by (2.30), controlled in closed-loop with the PI controller

$$C(s) = 0.1 + \frac{0.3}{s} \tag{2.36}$$

$B(s)$, given by (2.31), controlled in closed-loop with $C(s)$, is unstable. When a Smith predictor is used instead, stability is regained, as also seen in Figure 2.16.

Remark 2.5. It should be stressed that the Smith predictor does *not* eliminate the effect of the delay. The response of the controlled system will always be delayed. What the Smith predictor does is to eliminate the effect of the delay in the stability of the closed-loop — and this is already much.

2.8 Exercises

1. Use Octave to define a time vector and a unit-slope ramp as follows:

   ```
   t = 0 :  0.1 :  20;
   input = t;
   ```

   Then use command `lsim` to find the responses to this ramp for all the transfer functions in sections 2.3.1 and 2.3.2.

2. Use (2.1) to explain why it is that

   (a) $\frac{5}{s+10}$ is stable,

   (b) $\frac{5}{s-10}$ is unstable,

   (c) $\frac{s-10}{s+20}$ is stable,

   (d) $\frac{5}{s^2-4-12}$ is unstable,

   (e) $\frac{13}{2s^3+17s^2+47s+42}$ is stable and its step response has no oscillations,

   (f) $\frac{13}{9s^3+390s^2+1514s+12560}$ is stable and its step response has oscillations.

   Hint: you can find the roots of the polynomial in the denominator of the transfer function of (e) with Octave, using command

   ```
   roots([2 17 47 42])
   ```

3. Use Octave to find the step response of this transfer function, that models a WEC:

   $$\frac{Y(s)}{F(s)} = \frac{5 \times 10^{-6}}{s^2+3s+2}. \tag{2.37}$$

   Use the corresponding Ziegler-Nichols rules to find a P controller, a PI controller, and a PID controller for the plant.

4. Do the same for a WEC described by the following transfer function, that includes a delay:

   $$G_p(s) = \frac{8 \times 10^{-6}}{s+3}\, e^{-0.7s} \tag{2.38}$$

   Hint: Octave does not implement delays, but what is the effect of the delay?

5. Figure 2.17 is the step response of this transfer function, that may describe some electromechanical actuator:

   $$G_a(s) = \frac{-s+30}{s^2+20s+200} \tag{2.39}$$

   Use the corresponding Ziegler-Nichols rules to find a P controller, a PI controller, and a PID controller for this plant.

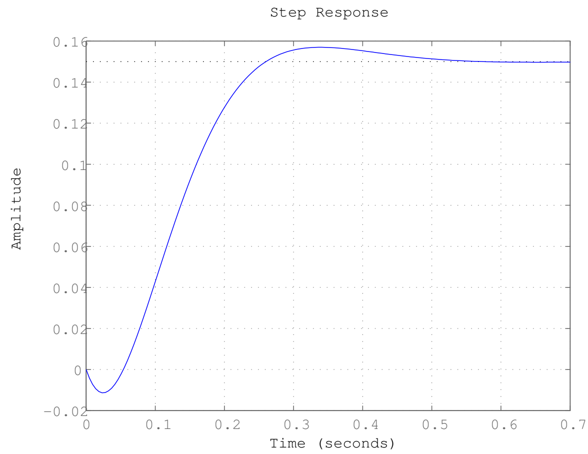

Figure 2.17: Step response of 2.39.

6. Suppose that WEC $G_p(s)$ from (2.38) is controlled with a proportional controller k, using a Smith predictor.

   (a) Find the transfer function of the closed loop, supposing that we have a perfect model of the delay ($\hat{\theta} = 0.7$) and a perfect model of the plant without delay ($G^* = \frac{8\times10^{-6}}{s+3}$).

   (b) Find the transfer function of the closed loop, supposing now that our model of the plant without delay is $G^* = \frac{8\times10^{-6}}{s+2}$.

7. Consider the control system in Figure 2.12. To simplify things, we will assume an almost perfect world: there are no disturbances, and the sensors give exact measurements instantaneously ($H_1(s) = H_2(s) = 1$).

   (a) How is this control structure called?

   (b) Which part of the plant do you expect to have a faster response: $G_1(s)$ or $G_2(s)$?

   (c) Find the transfer function of the inner closed loop, $\frac{Y_2(s)}{U_1(s)}$.

   (d) Find the transfer function of the entire system, $\frac{Y_1(s)}{R(s)}$.

   (e) Suppose that $G_2(s)$ is $G_a(s)$ given by (2.39), that $C_2(s)$ is the proportional controller you found for that plant, that $G_1(s)$ is $\frac{Y(s)}{F(s)}$ given by (2.37), and that $C_1(s)$ is the PID controller you found for that plant. Replace values in $\frac{Y_1(s)}{R(s)}$ and find the transfer function of the control system in that case.

8. Explain why on-off control is non-linear.

Chapter 3

Applications

This section introduces three papers that concern the control of two WECs and cover most of the subjects addressed so far.

3.1 The Archimedes Wave Swing (AWS)

The Archimedes Wave Swing was an off-shore submerged heaving point absorber, using a linear electric generator as PTO, and deployed in the north of Portugal (Figure 3.1).

[1] addresses the following issues:

- identification of a transfer function model;
- reactive control;
- phase and amplitude control (including proportional control);
- latching;
- feedback linearisation.

[2] addresses the following issues:

- identification of a NN model;
- IMC;
- switching control.

3.2 The Inertial Sea Wave Energy Converter (ISWEC)

The Inertial Sea Wave Energy Converter is an off-shore floating pitching point absorber, using gyroscopes and an electric generator as PTO, and deployed off Pantelleria island, Italy (Figure 3.2)

[3] addresses the following issues:

- PID control;
- switching control;

Figure 3.1: The Archimedes Wave Swing.

Figure 3.2: The Inertial Sea Wave Energy Converter.

- optimal control;
- nonlinear control.

Bibliography

In case you ever need to study these subjects in greater depth, you may wish to consult one or more of the following textbooks. [**?**] is a classical book, covering the control of continuous systems. [**?**], by the same author, covers discrete systems. [**?**] covers NNs and fuzzy modelling and control.

* * *

[1] Duarte Valério, Pedro Beirão, and José Sá da Costa. Optimisation of wave energy extraction with the Archimedes Wave Swing. *Ocean Engineering*, 34(17–18):2330–2344, 2007.

[2] Duarte Valério, Mário J. G. C. Mendes, Pedro Beirão, and José Sá da Costa. Identification and control of the AWS using Neural Network models. *Applied Ocean Research*, 30(3):178–188, 2008.

[3] Giacomo Vissio, Duarte Valério, Giovanni Bracco, Pedro Beirão, Nicola Pozzi, and Giuliana Mattiazzo. ISWEC Linear Quadratic Regulator oscillating control. *Renewable Energy*, 103:372–382, 2017.